

El Joc de les Bitlles



Les xapes

Material: xapes d'ampolla, plastilina i foto d'esportista preferit:
 Agafeu una xapa que no estigui doblegada, ompliu-la de plastilina i tapeu-la amb una foto



Feu una línia al terra a un metre de la paret. El joc consisteix a fer lliscar, amb un cop de dit, una xapa des d'una distància de 3 metres de la línia i que la xapa quedi el més prop possible. Es pot fer rebotar la xapa a la paret o no.

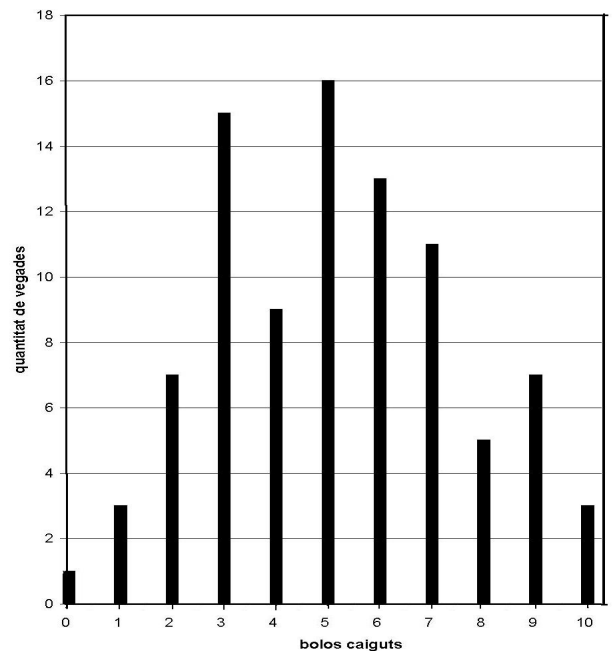
Organització de les dades en la variable discreta

Els resultats d'un equip en el joc de les bitlles són els següents:

2, 5, 9, 7, 3, 5, 10, 1, 4, 3, 9, 5, 4, 5, 7, 3, 10, 0, 7, 2, 3, 3, 3, 7, 5, 6, 4, 3, 2, 6, 9, 9, 5, 5, 3, 6, 9, 6, 6, 8, 2, 2, 5, 3, 1, 5, 4, 3, 4, 3, 4, 6, 8, 4, 7, 4, 7, 8, 2, 6, 8, 6, 7, 6, 5, 10, 5, 5, 6, 2, 3, 7, 3, 7, 1, 6, 5, 4, 5, 5, 6, 9, 3, 3, 8, 9, 4, 5, 6, 7, 7

El primer que hem fet amb els resultats d'un equip ha estat agrupar les dades en una taula i fer-ne un gràfic:

Número de bitlles tirades	Quantitat de vegades que les has tirades
0	1
1	3
2	7
3	15
4	9
5	16
6	13
7	11
8	5
9	7
10	3
Σ=Total tirades =	90



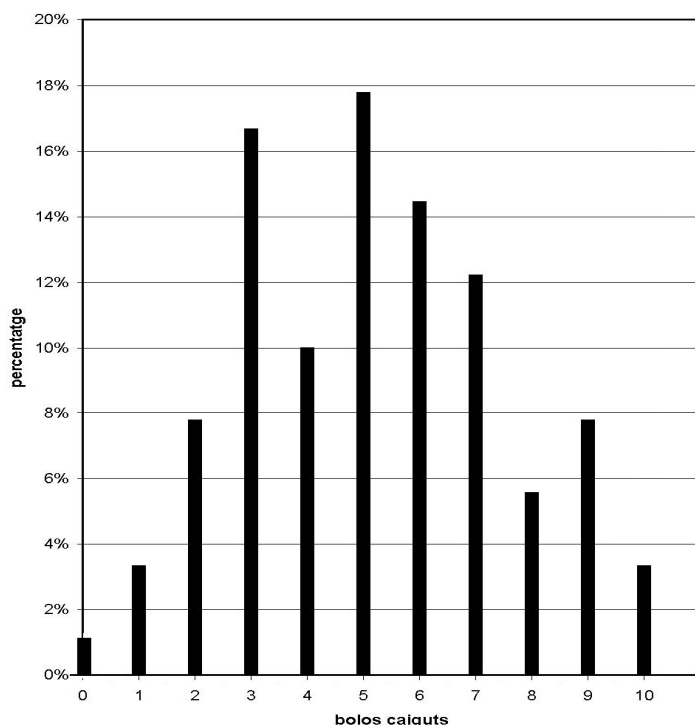
Càlculs amb percentatge.

El recompte de cada un dels valors de l'estudi estadístic rep el nom de **freqüència absoluta**.

Bitlles tirades	Freqüència absoluta	Freqüència relativa
0	1	1,11%
1	3	3,33%
2	7	7,78%
3	15	16,67%
4	9	10%
5	16	17,78%
6	13	14,44%
7	11	12,22%
8	5	5,56%
9	7	7,78%
10	3	3,33%
Σ=Total tirades =	90	100%

← $\frac{1}{90} = 0,0111 = \frac{1,11}{100} = 1,11\%$

← La resta de percentatges es calculen d'una manera similar



Organització de les dades en la variable contínua

Hem donat a un equip les xapes i les han estat tirant durant una estona. Hem estat mesurant la distància, en centímetres, de les xapes a la paret. Els resultats que han obtingut han estat:

138; 46; 184; 105; 84; 65; 68; 40; 41; 36; 170; 186; 87; 118; 92; 95; 71; 125; 59; 91; 21; 126; 105; 68; 126; 197; 116; 82; 59; 141; 155; 125; 2; 115; 85; 195; 55; 73; 17; 141; 63; 43; 115; 27; 68; 119; 78; 24; 118; 89; 116; 69; 32; 75; 16; 65; 71; 115; 106; 75.

Observa com, ara, a diferència del cas anterior de les bitlles, els valors poden ser qualsevol nombre entre 0 i 200, i per tant no podem fer una taula amb tots i cada un dels valors. En aquest cas hem d'agrupar els valors en grups (en l'estadística aquests grups s'anomenen **classes**). Podríem agrupar-los de 10 en 10 però tindríem 20 *classes* i son massa. Si els agrupem, però, de 20 en 20 en tindríem 10 que és molt més raonable:

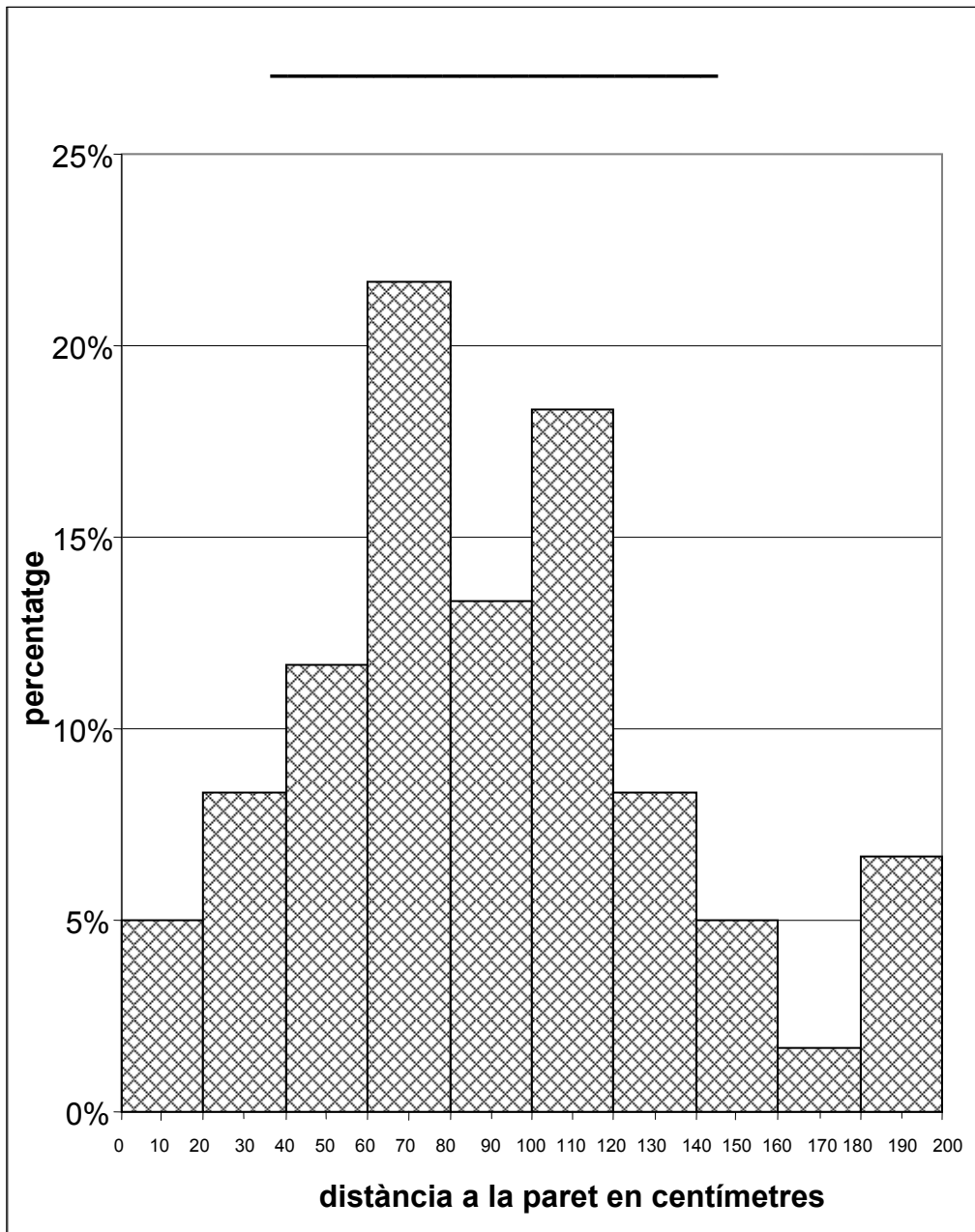
Distància a la paret	Freqüència Absoluta	Freqüència relativa
[0,20)	3	5 %
[20,40)	5	8,3 %
[40,60)	7	11,7 %
[60,80)	13	21,7 %
[80,100)	8	13,3 %
[100,120)	11	18,3 %
[120,140)	5	8,3 %
[140,160)	3	5 %
[160,180)	1	1,7 %
[180,200)	4	6,7 %
Total tirades =	60	100 %

A.1 Observa que els parèntesis utilitzats per expressar les diferents **classes** són diferents: un claudàtor al principi i un parèntesis ordinari al final: [20, 40). Recorda els intervals.

A.2 A la classe [40, 60) hi ha una freqüència absoluta de 7. Què vol dir això?. Com ha estat calculat aquest 7?

A.3 La freqüència relativa del mateix interval [40, 60) és 11,7%? Què volia dir la freqüència relativa?. Com ha estat calculat el 11,7 %?

La taula d'aquest exemple és diferent a la de l'exemple anterior. Ara les dades estan separades en intervals o classes. El gràfic haurà de reflectir, d'alguna manera, aquesta diferència i haurà de ser diferent. Observa el gràfic:



Posem noms a les coses

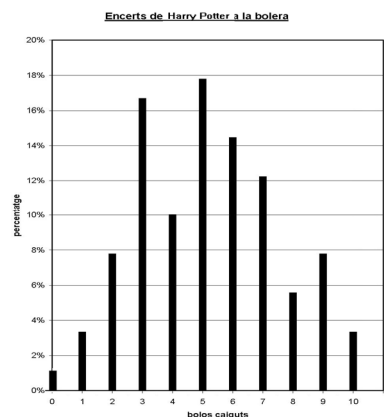
Ja hem treballat una quantitat raonable de conceptes. Abans de continuar cal que tinguem clar com s'anomena cada una de les coses que fem per tal que ens aclarim quan parlem o escrivim les nostres idees. Algunes de les coses ja han sortit però ara les definirem d'una manera més estricta.

- Anomenem **població** el conjunt d'individus sobre els quals es fa un determinat estudi.
- Anomenem **caràcter** o **variable** a l'aspecte de la població que s'estudia. Per exemple el caràcter que estudiem en el primer exemple és la quantitat de bitlles que poden tirar jugant al joc de bitlles

Tipus de caràcters

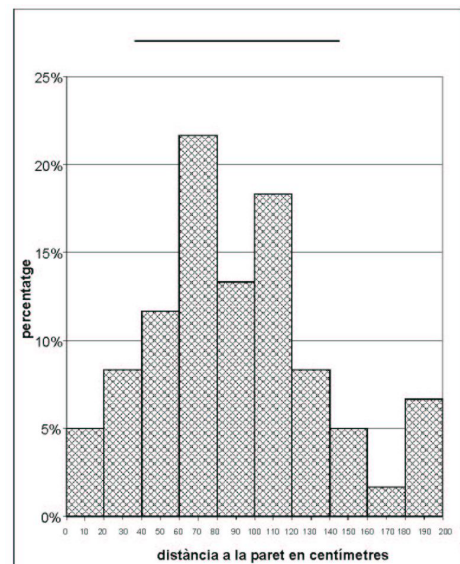
Observa com els dos exercicis anteriors són molt diferents:

- En el cas del joc de bitlles de cap manera podem fer caure 3,7 bitlles, on en fem caure 3 o 4 però no hi ha cap possibilitat de tirar-ne cap entre el 3 i el 4. A aquest tipus de caràcter l'anomenem **variable discreta** i el gràfic que requereix és un **diagrama de barres**. En el que les barres han d'estar separades. A l'exemple de les bitlles la barra neix del 3, o del 4 però no hi ha res entre el 3 i el 4.



- Les xapes, per contra, poden estar a qualsevol distància de la paret i entre 20 i 40 centímetres, per exemple, les xapes poden prendre qualsevol valor. Com que tots els valors són possibles no hi pot haver separació entre l'agrupació de 20 a 40 i la següent de 40 a 60.

Aquest tipus de caràcters reben el nom de **variable contínua** i el gràfic que requereixen s'anomena **histograma**. En un **histograma** els rectangles han d'estar junts per què, de fet, tots els valors són possibles, i per tant el rectangle ha d'abastar tots els valors que representa l'interval. A l'exemple el rectangle de la classe [20, 40) comença realment al nombre 20 i acaba al 40, però la classe [40, 60) comença també a 40. Per tant, per força s'ha de tocar amb el rectangle de l'interval anterior.



Els quartils i els diagrames de Caixa (Box plot)

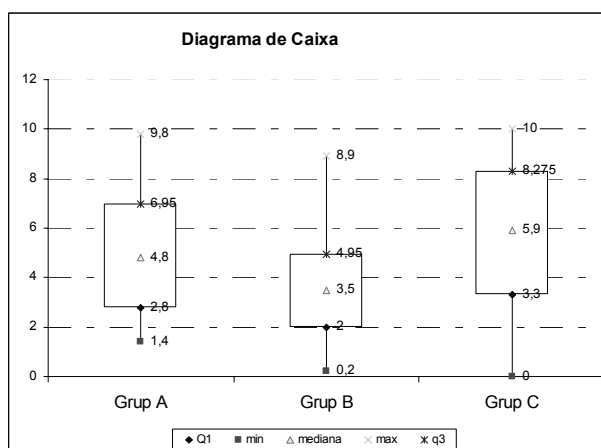
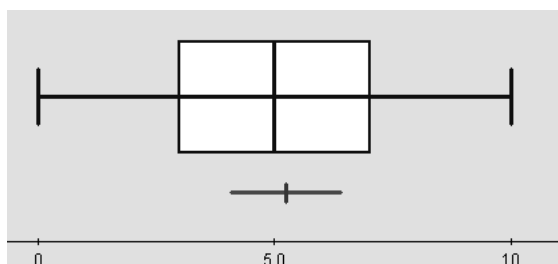
Si la mediana ens indica en quina posició trobem el 50% de les dades, els quartils ens indiquen on es troben el 25% i el 75% de les dades. Així, el Q1 és el primer quartil, Q2 la mediana i en canvi, Q3 és el tercer quartil.

En la taula anterior hi podem localitzar els quartils. De 90 tirades, hem dit que la meitat, el 50% són 45, els primer i tercer quartil se situen a les tirades 23 i 68. Així: Q1=3, Q3=7.

Número de bitlles tirades	Quantitat de vegades que les has tirades	Tirades acumulades	
0	1	1	
1	3	1+3 = 4	
2	7	4+7 = 11	
3	15	11 + 15 = 26	← Q1 és aquí
4	9	26 + 9 =35	
5	16	35 + 16 = 51	← Q2 és aquí
6	13	51 + 13 = 64	
7	11	64 + 11 = 75	← Q3 és aquí
8	5		
9	7		
10	3		
Total tirades =	90		

El diagrama de caixa mostra les dades extremes i les quartils d'un conjunt de dades i en mostra la posició relativa. Per representar un diagrama de caixa, es dibuixa un rectangle de base igual al segment que separa el Q1 i el Q3, partit en dos per una línia que té el valor de la mediana. A banda i banda del rectangle es dibuixen uns segments, anomenats cues que representen tot el rang de dades.

Finalment, un eix graduat segons els valors que pot tenir la variable permetrà entendre el gràfic: Sovint pot convenir comparar en un mateix gràfic diversos diagrames de caixa



La mitjana

Si ho penses bé, veuràs com els paràmetres anteriors tenen el mateix defecte: en el seu càlcul no es tenen en compte, de fet, totes les dades

Anomenem **mitjana** a la suma de tots els resultats dividit entre la quantitat de resultats. La mitjana l'escriurem amb la lletra \bar{x} .

Sumem les bitlles caigudes:

2 + 5 + 9 + 7 + 3 + 5 + 10 + 1 + 4 + 3 + 9 + 5 + 4 + 5 + 7 + 3
 + 10 + 0 + 7 + 2 + 3 + 3 + 3 + 7 + 5 + 6 + 4 + 3 + 2 + 6 + 9 + 9 +
 5 + 5 + 3 + 6 + 9 + 6 + 6 + 8 + 2 + 2 + 5 + 3 + 1 + 5 + 4 + 3 + 4
 + 3 + 4 + 6 + 8 + 4 + 7 + 4 + 7 + 8 + 2 + 6 + 8 + 6 + 7 + 6 + 5 +
 10 + 5 + 5 + 6 + 2 + 3 + 7 + 3 + 7 + 1 + 6 + 5 + 4 + 5 + 5 + 6 + 9
 + 3 + 3 + 8 + 9 + 4 + 5 + 6 + 7 + 7 = ?

Millor els ordenem:

0 + 1 + 1 + 1 + 2 + 2 + 2 + 2 + 2 + 2 + 2 + 3 + 3 + 3 + 3 + 3 + 3 + 3
 + 3 + 3 + 3 + 3 + 3 + 3 + 3 + 3 + 4 + 4 + 4 + 4 + 4 + 4 + 4 + 4 +
 5 + 5 + 5 + 5 + 5 + 5 + 5 + 5 + 5 + 5 + 5 + 5 + 5 + 5 + 5 + 6 + 6
 + 6 + 6 + 6 + 6 + 6 + 6 + 6 + 6 + 6 + 6 + 7 + 7 + 7 + 7 + 7 +
 7 + 7 + 7 + 7 + 8 + 8 + 8 + 8 + 8 + 8 + 9 + 9 + 9 + 9 + 9 + 9 +
 10 + 10 + 10 = ?

Per calcular això amb la calculadora és un embolic i és fàcil d'equivocar-nos.

Però encara és molt més fàcil si multipliquem primer aquells que estan repetits:

$$0 \times 1 + 1 \times 3 + 2 \times 7 + 3 \times 15 + 4 \times 9 + 5 \times 16 + 6 \times 13 + 7 \times 11 + 8 \times 5 + 9 \times 7 + 10 \times 3 = ?$$

Però aquest càlcul és molt més senzill si el fem directament a la taula afegint una columna:

Número de bitlles tirades	Quantitat de vegades que les has tirades	(Bitlles tirades)X (vegades)
0	1	0x1 = 0
1	3	1x3 = 3
2	7	2x7 = 14
3	15	3x15 = 45
4	9	4x9 = 36
5	16	5x16 = 80
6	13	6x13 = 78
7	11	7x11 = 77
8	5	8x5 = 40
9	7	9x7 = 63
10	3	10x3 = 30
Suma =	90	466

L'equip ha tirat un total de 466 bitlles en 90 tirades. Li toquen a:

$$\bar{x} = \frac{466}{90} = 5,18 \text{ bitlles tirades per cada llançament fet.}$$

Paràmetres de centralització a la variable contínua.

La majoria de vegades no es disposa de les dades originals i únicament tenim les dades amb una taula del tipus següent:

Distància a la paret	Freqüència Absoluta	Freqüència relativa	Freq. relativa acumulada	
[0,20)	3	5 %	5%	
[20,40)	5	8,3 %	8,3% + 5% = 13,3%	
[40,60)	7	11,7 %	11,7% + 13,3 % = 25%	← Q1 és aquí
[60,80)	13	21,7 %	21,7% + 25% = 46,7 %	
[80,100)	8	13,3 %	13,3% + 46,7% = 60%	← Q2 és aquí
[100,120)	11	18,3 %	18,3% + 60% = 78,3 %	← Q3 és aquí
[120,140)	5	8,3 %	8,3 % + 78,3 % = 86,6 %	
[140,160)	3	5 %	5% + 86,6% = 91,6 %	
[160,180)	1	1,7 %	1,7 % + 91,6% = 93,3 %	
[180,200)	4	6,7 %	6,7 % + 93,3 = 100%	
Total tirades =	60	100 %		

En aquest cas no podem saber exactament quin és el mode, la mediana o els quartils, però podem dir que.

- El **mode** és entre 60 i 80. o La **classe modal** és [60,80), Ja que aquesta classe és la que té més dades (la més repetida)
- La **mediana** és entre 80 i 100 o la **classe mediana** és [80,100) ja que el valor del mig és el que fa 30 i si anem sumant la freqüència absoluta veiem que el que fa 30 és dins del interval [80,100)
- Indica quins són **Q1** i **Q3**.

La **mitjana** no es pot calcular amb exactitud però podem fer un càlcul aproximat imaginant que els tres cops que la xapa ha estat entre 0 i 20 ha quedat exactament al mig: a 10 cm, els 5 cops que ha estat entre 20 i 40 imaginem que ha estat exactament a 30 etc:

Anomenem **marca** de la classe al valor que està al mig de la classe.

Per calcular la **mitjana** imaginarem totes les vegades que han sortit els valors d'una classe en realitat han sortit els de la marca (el valor del mig)

Observa l'exemple de les xapes:

Distància a la paret	Marca	Freqüència Absoluta	Marca x Freq Abs.
[0,20)	10	3	10 x 3 = 30
[20,40)	30	5	30 x 5 = 150
[40,60)	50	7	50 x 7 = 350
[60,80)	70	13	70 x 13 = 910
[80,100)	90	8	90 x 8 = 720
[100,120)	110	11	110 x 11 = 1210
[120,140)	130	5	130 x 5 = 650
[140,160)	150	3	150 x 3 = 450
[160,180)	170	1	170 x 1 = 170
[180,200)	190	4	190 x 4 = 740
Suma =		60	5380

En total, si sumem la distància de totes les xapes a la paret suma, aproximadament 5380 que dividit entre 60 vegades que hem tirat la xapa, toca a: $\bar{x} = \frac{5380}{60} = 89,67$ cm de distància mitjana per cada xapa.

Càlcul de paràmetres estadístics:

Mitjana:
$$\bar{x} = \frac{\sum_{i=1}^m x_i \cdot n_i}{n}; \quad n = \sum_{i=1}^m n_i$$

Variància:
$$\sigma^2 = \frac{\sum (x_i - \bar{x})^2 \cdot n_i}{n} = \sum \frac{x_i^2 \cdot n_i}{n} - \bar{x}^2$$

Desviació Típica:
$$\sigma = \sqrt{\frac{\sum (x_i - \bar{x})^2 \cdot n_i}{n}} = \sqrt{\sum \frac{x_i^2 \cdot n_i}{n} - \bar{x}^2}$$

Com es calcula la desviació típica?

La calculadora per fer les operacions ha de conèixer quines són les dades del problema. La primera cosa a fer és, doncs, introduir les dades. Començarem posant la calculadora al MODE ESTADÍSTIC. A la majoria de calculadores s'ha de seguir la seqüència **MODE** **▢** ; a altres s'ha de pitjar la tecla **MODE** i es busca l'opció SD dins del menú que apareix. El més important és que han d'aparèixer a la pantalla de la teva calculadora les lletres SD. Quan ja no treballes amb l'estadística has de fer que la calculadora torni al MODE COMP que, normalment, es fa teclejant **MODE** **1** o fent ús del menú que surt al pitjar la tecla **MODE**



A.4 Fes proves a la teva calculadora. Fixa't com apareix i desapareix de la pantalla el símbol SD. Escribeu la seqüència de la teva calculadora per posar-la en MODE SD i per a tornar-la al MODE COMP.

Pregunta a respondre	seqüència de tecles dependent del model
Passar d'un mode SD a COMP i viceversa	

Posa la teva calculadora amb el MODE SD. Quan tenim el MODE SD a la pantalla, la calculadora està preparada per fer els càlculs estadístics fent servir les lletres de color blau que hi ha sota les tecles dels nombres i d'algunes altres tecles.

Primer hem d'estar segurs que a la memòria no hi ha res que pugui interferir als nostres càlculs: farem servir la seqüència **SHIFT** **AC** (fixa't que en blau, sobre la tecla **AC**, posa SAC que són les inicials de "Statistic All Clear", (esborrar totes les dades estadístiques). A altres cal teclejar **SHIFT** **MODE** i amb el menú que surt, triar l'opció SCL.

Procedim a introduir les dades una a una teclejant després de cada dada la tecla **M+**. Continua escrivint la seqüència per introduir les dades: Per exemple, si volem introduir les dades 1,10; 1,15 i 1,05 premerem:

1.10 **M+** 1.15 **M+** 1.05 **M+** ...

Si hem d'introduir una dada més d'una vegada (per exemple, la dada 1.10, tres vegades) podem fer servir la seqüència:

1.10 **X** 3 **M+** o dependent del model 1.10 **SHIFT** **;** 3 **M+**

Una vegada introduïdes les dades ja podem calcular els diversos paràmetres de centralització, de dispersió i fins i tot algunes coses més.

Pregunta a respondre	seqüència de tecles dependent del model	Símbol	Resultat amb la calculadora
Quantes dades hem introduït?	SHIFT 6 RCL hyp SHIFT 1 3 =	n	
Quina és la suma total de totes les dades?	SHIFT 5 RCL 0 = SHIFT 1 2 =	Σx	
Quina és la suma total de les dades cadascuna elevada al quadrat?	SHIFT 4 RCL (-) SHIFT 1 1 =	Σx^2	
Quina és la mitjana aritmètica d'aquestes dades?	SHIFT 7 SHIFT 1 = SHIFT 2 1 =	\bar{x}	
Quina és la quasi desviació típica dels valors respecte de la mitjana?	SHIFT 9 SHIFT 3 = SHIFT 2 3 =	σ_{n-1}	
Quina és la desviació típica dels valors respecte de la mitjana?	SHIFT 8 SHIFT 2 = SHIFT 2 2 =	σ_n	

Alguns models de calculadores permeten corregir les possibles errades quan entrem dades. Si la teva calculadora disposa de les tecles de cursor, utilitza-les per visualitzar els valors $x_j = i$ *Freqj*= que pots modificar.

Càlcul de la desviació tipus a mà

Per calcular la desviació tipus a mà cal omplir la taula següent. A la fila de sota cal fer la suma de tots els valors de la columna corresponent. El símbol que s'utilitza per expressar aquesta suma és la lletra grega "sigma" Σ .

Dades o marca de la classe	Frequència absoluta			
x_i	n_i	$x_i \cdot n_i$	x_i^2	$x_i^2 \cdot n_i$
x_1	n_1	$x_1 \cdot n_1$	x_1^2	$x_1^2 \cdot n_1$
x_2	n_2	$x_2 \cdot n_2$	x_2^2	$x_2^2 \cdot n_2$
.
.
x_k	n_k	$x_k \cdot n_k$	x_k^2	$x_k^2 \cdot n_k$
Suma =	N	$\sum_{i=1}^k x_i n_i$		$\sum_{i=1}^k x_i^2 n_i$

La **quantitat total de dades** és $N = \sum_{i=1}^n n_i$ i la fórmula per calcular la **desviació**

tipus és
$$\sigma = \sqrt{\frac{\sum_{i=1}^k x_i^2 n_i}{N} - \bar{x}^2}$$

Exemple: L'Andreu ha fet un total de 20 exàmens de història durant el curs. Amb els següents resultats: 6 cops ha tret un 4, 8 cops un 5, 4 cops un 6 i 2 cops un 7. La taula que descriu la seva situació estadística de l'assignatura és

Notes x_i	Quantitat n_i	$x_i \cdot n_i$	x_i^2	$x_i^2 \cdot n_i$
4	6	24	16	96
5	8	40	25	200
6	4	24	36	144
7	2	14	49	98
Suma =	20	102		538

La **quantitat total de dades** és $N = \sum_{i=1}^n n_i = 20$

La **mitjana** és calcula fent $\bar{x} = \frac{\sum_{i=1}^k x_i n_i}{N} = \frac{102}{20} = 5,1$

La fórmula per calcular la **desviació tipus** és

$$\sigma = \sqrt{\frac{\sum_{i=1}^k x_i^2 n_i}{N} - \bar{x}^2} = \sqrt{\frac{538}{20} - 5,1^2} = \sqrt{26,9 - 26,01} = \sqrt{0,89} \approx 0,94$$

**Explica com trobes la mitjana i la desviació tipus amb la teva calculadora.
Quin model és?
Quina és la seqüència de tecles?**

Aplicació:

A.5 L'alçada i l'edat, inventades, dels jugadors de Los Angeles Lakers i dels Chicago Bulls de la NBA les tens en aquestes taules.

L.A.L.	edat	23	25	28	32	30	30	30	31	30	30	27	22
	alçada	2.10	2.15	2.05	2.05	2.06	1.93	1.88	1.95	2.07	1.87	2.08	1.94

CH.B.	edat	25	34	26	31	26	28	30	31	26	26	24	22
	alçada	1.87	2.15	2.08	1.88	1.95	1.98	2.03	1.88	2.13	2.00	2.10	2.05

- Calcula la mitjana i la desviació tipus de l'alçada de cada equip.
- Quin equip és més alt de mitjana?
- Calcula la mitjana i desviació tipus de l'edat de cada equip.
- Quin equip és més jove de mitjana?
- Fes un diagrama de caixa que compari les edats dels dos equips i un altre que compari les alçades dels dos equips.

A.6 Les qualificacions obtingudes per un grup d'alumnes en l'assignatura de ciències de la naturalesa són les següents:

8,8 3,1 4,7 6,3 4,5 9,6 8,8 8,3 3,3 7,2
2,0 8,4 4,4 8,7 5,6 2,2 7,3 2,2 7,6 5,5
6,3 4,5 3,8 7,5 5,4 8,0 3,9 4,6 8,4 6,8
5,3 3,4 4,3 2,5 7,2 5,0 3,6 5,6 7,2 9,2
1,8 4,3 5,0 5,4 5,5 5,8 4,0 5,9 8,2 7,6

- Agrupa les dades en les classes $[0, 2)$; $[2, 4)$; $[4, 6)$; $[6, 8)$; $[8, 10)$ i calcula'n les freqüències absolutes, les relatives i les freq. absolutes acumulades i les freq. relatives acumulades.
- Calcula la classe modal i la classe mediana.
- Fes un histograma.
- Calcula la mitjana i la desviació tipus. Mostra els càlculs i els passos que has fet.